REMARKS

Applicant requests favorable reconsideration and allowance of the subject application based on the following remarks. Claims 1, 3-6, 10-19, and 28-44 remain pending in the application and are for reconsideration. No new matter has been introduced by this response.

Applicant's response and remarks after Final are appropriate under 37 C.F.R. §1.116 because they address the Office's remarks in the Final Action, and thus could not have been presented earlier. In addition, the remarks should be entered to place the application in better form for appeal.

Inconsistency in Claim Rejections

There have been six office actions from the Examiner regarding this application dating from March 2004 to November 2006. Applicant in good faith has attempted to advance the prosecution for this application. In particular, Applicant and other representatives have amended the claims, but all our efforts availed us little in advancing the prosecution. Thus, Applicant respectfully points out several inconsistencies in the Office's continuing rejections for this application.

Claim 7 as amended into Claim 1

Applicant submits that the claim rejections are not in accordance with the MPEP, 35 U.S.C. §102, and §103. First, the Office rejects **Claim 7** under §103(a) for obviousness over U.S. Patent No. 6,188,976 to Ramaswamy, in view of U.S. Patent No. 6,317,707 to Bangalore. The Office relies on Bangalore as providing the teachings of

Claim 7 (dated 3/25/04 on pages 6-7). Applicant amends independent Claim 1 to recite the features along the lines of dependent Claim 7 (dated 5/24/04). Accordingly, Claim 7 is cancelled in the same response.

Next, the Office rejects amended Claim 1 in two more office actions under §103 for obviousness citing Ramaswamy in view of Bangalore. As before, the Office relies on Bangalore providing the teachings of former Claim 7 (Office Action of 8/11/04 on page 4 and 2/28/05 on pages 3-4). Thus, the Office cites Bangalore three times.

Applicant points out the first inconsistency when the Office rejects amended Claim 1 under §102(e) for anticipation by Ramaswamy in two separate office actions (11/16/05 on pages 3-4 and 4/19/06 on pages 2-3). Applicant finds perplexing how the features of Claim 7 as amended into Claim 1, were initially rejected for obviousness over Ramaswamy in view of Bangalore in two previous office actions, but later rejected for anticipation by Ramaswamy in two other separate office actions. As previously noted, Bangalore was cited as providing the features of Claim 7 for the §103 rejection in three separate office actions, but is no longer cited. Thus, the Office may not assert Ramaswamy describes features that it does not disclose, teach or suggest, just because Applicant amended the claim. Also, the Office fails to provide any explanation changing the rejection of the features of former Claim 7 from a §103 rejection with two references to a §102 rejection with one reference.

The inconsistency continues when the Office changes the rejection on Claim 1 from a §102 to a §103 as being unpatentable over Ramaswamy in view of another reference, the publication, "N-th Order Ergodic Multigram HMM for Modeling of Languages without Marked Word Boundaries" by Law et al in Office Action, dated

11/08/2006 on pages 4-5. The Office had stated that Ramaswamy does not disclose, teach or suggest the features of Claim 7 (cited in three separate office actions). Therefore, the Office may not assert Ramaswamy describes features that are not disclosed, taught, or suggested, because Applicant amended the claim. Thus, the rejection of Claim 1 should be withdrawn.

Claim 9 as amended into Claim 1

Regarding Claim 9, Applicant respectfully points out similar inconsistencies in the Office's rejections. The Office rejects **Claim 9** under §103(a) for obviousness over Ramaswamy in view of Bangalore. Four Office Actions rely on Bangalore as providing the teachings of Claim 9 (Office Actions: of 3/25/04 on page 7, of 8/11/04 on page 7, of 2/28/05 on page 6, and of 11/16/2005 on page 11). Applicant amends independent Claim 1 to recite the features along the lines of dependent Claim 9 (2/9/06). Accordingly, Claim 9 is cancelled in the same response.

Applicant points out the inconsistency when the Office changes the rejection of amended Claim 1 from §103(a) to §102(e) as being allegedly anticipated by Ramaswamy (Office Action dated 4/19/06 on pages 2-3). Again, Applicant finds perplexing how the features of Claim 9 as amended into Claim 1, initially rejected for obviousness by Ramaswamy in view of Bangalore in four previous office actions, but later rejected for anticipation by Ramaswamy in one separate office action. Bangalore was cited as providing the features of Claim 9 for the §103 rejection in four office actions, but is no longer cited. Furthermore, the Office fails to provide any explanation changing the

rejection of the features of Claim 9 from a §103 rejection with two references to a §102 rejection with one reference.

Applicant indicates how the inconsistency continues as the Office changes the rejection of Claim 1 from a §102 to a §103 for obviousness over Ramaswamy in view of Law (Office Action dated 11/08/2006 on pages 4-5). The Office had stated that Ramaswamy does not disclose, teach or suggest the features of Claim 9. Therefore, Applicant submits the Office may not assert Ramaswamy describes features that are not disclosed, taught or suggested, just because Applicant amended the claim. Accordingly, this rejection should be withdrawn.

Claim Rejections 35 U.S.C. §103

A. Claims 1, 3-6, 14-19, 28-30, 33, and 35-39 stand rejected under 35 U.S.C. §103(a) as being unpatentable over U.S. Patent No. 6,188,976 to Ramaswamy et. al (hereinafter "Ramaswamy") in view of the publication, "N-th Order Ergodic Multigram HMM for Modeling of Languages without Marked Word Boundaries" by Law et. al (hereinafter "Law"). Applicant respectfully traverses the rejections.

Independent Claim 1 recites:

> *A method of using a tuning set of information to jointly optimize the performance and size of a language model, comprising:*
> *providing a textual corpus comprising subsets wherein each subset comprises a plurality of items;*
> *creating a Dynamic Order Markov Model data structure by assigning each item of the plurality of items to a node in the data structure, wherein the nodes are logically coupled to denote dependencies of the items, and calculating a frequency of occurrence for each item of the plurality of items;*
> *segmenting at least a subset of a received textual corpus into segments by clustering every N-items of the received corpus into a training unit, wherein resultant training units are separated by gaps, and*

wherein N is an empirically derived value based, at least in part, on the size of the received corpus;

   creating the tuning set from application-specific information;

   (a)training a seed model via the tuning set;

   (b)calculating a similarity within a sequence of the training units on either side of each of the gaps;

   (c)selecting segment boundaries that maximize intra-segment similarity and inter-segment disparity;

   (d)calculating a perplexity value for each segment based on a comparison with the seed model;

   (e)selecting some of the segments based on their respective perplexity values to augment the tuning set;

   iteratively refining the tuning set and the seed model by repeating parts (a) through (e) with respect to a threshold; and

   refining the language model based on the seed model.

References Fail to Teach or Suggest N is empirically..., calculating..., selecting segment:

Applicant agrees with the Office's assessment in several previous office actions that Ramaswamy does not disclose the features of former Claims 7 and 9, as amended into Claim 1. Therefore, Ramaswamy does not disclose, teach, or suggest *"wherein N is an empirically derived value based, at least in part, on the size of the received corpus* (former Claim 9), *calculating a similarity within a sequence of the training units on either side of each of the gaps; selecting segment boundaries that maximize intra-segment similarity and inter-segment disparity* (former Claim 7)*"*, as recited in Claim 1.

First, the current Office Action contradicts previous ones by citing portions of Ramaswamy describing a number n can either be a predetermined fixed number or a number that dynamically varies with each language model building iteration (col. 3, lines 50-52), which is <u>not</u> an empirically derived value based on the size of the received corpus, as recited in Claim 1. Second, the cited portions of Ramaswamy describe a relevance score calculator calculates the relevance score of each linguistic unit read from external

corpus (col. 6, lines 34-36). This is <u>not</u> disclosing calculating a similarity within a sequence of the training units on either side of each of the gaps, as recited in Claim 1. Third, Ramaswamy as cited, describes threshold calculator takes the relevance score of the currently read linguistic unit and places the linguistic unit in the appropriate subcorpus (col. 6, lines 36-41), which is <u>not</u> selecting segment boundaries that maximize intra-segment similarity and inter-segment disparity, as recited in Claim 1. Applicant asserts the evidence as cited fails to disclose, teach, or suggest *"wherein N is an empirically derived value based, at least in part, on the size of the received corpus* (former Claim 9), *calculating a similarity within a sequence of the training units on either side of each of the gaps; selecting segment boundaries that maximize intra-segment similarity and inter-segment disparity"* (former Claim 7), as recited in Claim 1.

If proposed modification would render the prior art invention being modified unsatisfactory for its intended purpose, then there is no suggestion or motivation to make the proposed modification (MPEP §2143.01 V.). Ramaswamy generates a domain-specific language model using a small, domain-restricted seed corpus, and a large, less restricted external corpus (col. 2, lines 45-47). Law is directed to N-th Order HMM for modeling of languages without word boundaries between adjacent words in a sentence for Oriental languages, such as Chinese (Abstract). In Law, the Ergodic Multigram Hidden Markov Model can process directly on unsegmented input corpus, a raw, untagged corpus (section 2.1 and 5). Thus, the proposed modification would render Ramaswamy and Law inoperable for their intended purposes, domain-specific language modeling and modeling of languages without word boundaries (Chinese, Japanese, Korean), respectively. Thus, the Office has failed to establish a *prima facie* case of obviousness. Accordingly, Claim 1

is allowable over Ramaswamy, Law, alone or in combination, for at least the reasons described above.

Bangalore was not cited for a §103 rejection for these claims in the current Office Action (11/8/06). However, Applicant points out that Bangalore does not disclose, teach or suggest *"wherein N is an empirically derived value based, at least in part, on the size of the received corpus"*, as recited in Claim 1. The Office had cited Bangalore as showing former Claim 9, which is reproduced below:

> *Bangalore, col. 2, lines 59-65:*
> Based upon the frequencies, **an N dimensional vector may be built for each input word.** The number of dimensions N of the frequency vector is a multiple of the total number of context words, the total number of input words and the total number of relations identified by the method 1000. The vector represents grammatical links that exist between the input words and the context words.

Applicant submits that Bangalore merely describes the creation of a N dimensional vector for each input word, where the vector represents grammatical links that exist between the input words and the context words. The Office appears to have included Bangalore simply because a variable "N" is derived from the total number of context words, the total number of input words and the total number of relations identified by the method 1000. Instead, the variable "N" used in Bangalore is not used for clustering every N-items of the received corpus into a training unit ... **wherein N is an empirically derived value based, at least in part, on the size of the received corpus.**

Consequently, Applicant respectfully submits that Claim 1 is allowable over Ramaswamy, Bangalore, or Law, and requests that the §103 rejection be withdrawn.

Dependent Claims 3-6 and 14-19 depend directly or indirectly from Claim 1, and thus are allowable as depending from an allowable base claim. These claims are also

allowable for their own recited features that, in combination with those recited in Claim 1, are not disclosed by Ramaswamy, Bangalore, or Law.

Independent Claims 28 and 36 recite features similar to those in Claim 1 and hence benefits from the same arguments directed above to Claim 1.

Dependent Claims 29-30, 33, 35, and 37-39 depend directly or indirectly from one of independent Claims 28 or 36, and thus are allowable as depending from an allowable base claim. These claims are also allowable for their own recited features that, in combination with those recited in Claims 28 and 36, are not disclosed, taught or suggested by Ramaswamy, Bangalore, or Law. Applicant respectfully submits that these claims are allowable over Ramaswamy, Bangalore, or Law, and requests that the §103 rejection be withdrawn.

B. Claims 10-13, 31 and 34 stand rejected under 35 U.S.C. §103(a) as being unpatentable over Ramaswamy in view of Law and further in view of Bangalore. Applicant respectfully traverses the rejections.

Dependent Claims 10-13, 31, and 34 depend directly or indirectly from one of independent Claims 1 and 28 and hence benefit from the same arguments directed above to Claim 1. These claims are allowable as Ramaswamy, Law, or Bangalore, alone or in combination, do not disclose, teach, or suggest *"N is an empirically derived value based, at least in part, on the size of the received corpus* (former Claim 9), *calculating a similarity within a sequence of the training units on either side of each of the gaps; selecting segment boundaries that maximize intra-segment similarity and inter-segment disparity"* (former Claim 7), as recited in Claim 1.

As previously mentioned, Ramaswamy generates a domain-specific language model using a small, domain-restricted seed corpus, and a large, less restricted external corpus (col. 2, lines 45-47). Law is directed to N-th Order HMM for modeling of languages without word boundaries between adjacent words in a sentence for Oriental languages, such as Chinese (Abstract), processing directly on unsegmented input corpus, a raw, untagged corpus (section 2.1 and 5). Thus, the proposed modification would render Ramaswamy and Law inoperable for their intended purposes, domain-specific language modeling and modeling of languages without word boundaries (Chinese, Japanese, Korean), respectively. Thus, the Office has failed to establish a *prima facie* case of obviousness.

The Office has failed to establish a *prima facie* case of obviousness. There must be some suggestion or motivation, either in the references themselves or in the knowledge generally available to one of ordinary skill in the art, to modify the reference or to combine reference teachings (MPEP §2142). The Office has failed to establish a motivation sufficient for one of ordinary skill in the art to modify the reference or to combine the references. The motivation provided by the Office to "determine how close or similar the training units were to each other for the benefit of maximizing the clustering process of related items" is not specific, as there is no suggestion or motivation by the references. The Office cannot improperly rely on hindsight without evidence of motivation to propose the suggested combination. Applicant respectfully requests the §103 rejection be withdrawn.

Since there is no suggestion or motivation, and the features are not disclosed, taught, or suggested in Ramaswamy, Law, or Bangalore alone, the resultant combination

of these references does not result in Applicant's Claims 10-13, 31 and 34. Accordingly, these claims are allowable over these references, individually or in combination for at least these reasons. Applicant respectfully requests that the §103 rejection of these claims be withdrawn.
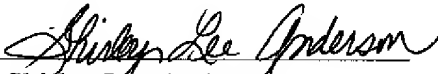
Conclusion

The pending Claims 1, 3-6, 10-19, and 28-44 are in condition for allowance. Applicant respectfully requests reconsideration and prompt issuance of the present application. If any issues remain that preclude issuance of the application, the Examiner is urged to contact the undersigned attorney before issuing a subsequent Action.

Respectfully Submitted,

Lee & Hayes, PLLC
421 W. Riverside Avenue, Suite 500
Spokane, WA 99201

Dated: _01-05-2007_   By: _Shirley Lee Anderson_

Shirley Lee Anderson
Reg. No. 57,763
(509) 324-9256 ext. 258